Hypertension Research (2009) 32, 1119–1122

www.nature.com/hr

ORIGINAL ARTICLE

Development and validation of a cardiovascular risk prediction model for Japanese: the Hisayama study

Hisatomi Arima[1], Koji Yonemoto[1], Yasufumi Doi[2], Toshiharu Ninomiya[1], Jun Hata[1], Yumihiro Tanizaki[1], Masayo Fukuhara[1], Kiyoshi Matsumura[2], Mitsuo Iida[2] and Yutaka Kiyohara[1]

The objective of this paper is to develop a new risk prediction model of cardiovascular disease and to validate its performance in a general population of Japanese. The Hisayama study is a population-based prospective cohort study. A total of 2634 participants aged 40 years or older were followed up for 14 years for incident cardiovascular disease (stroke and coronary heart disease (myocardial infarction, coronary revascularization and sudden cardiac death)). We used data among a random two-thirds (the derivation cohort, *n*=1756) to develop a new risk prediction model that was then tested to compare observed and predicted outcomes in the remaining one-third (the validation cohort, *n*=878). A multivariable cardiovascular risk prediction model was developed that incorporated age, sex, systolic blood pressure, diabetes, low-density lipoprotein cholesterol, high-density lipoprotein cholesterol and smoking. We assessed the performance of the model for predicting individual cardiovascular event among the validation cohort. The risk prediction model demonstrated good discrimination (c-statistic=0.81; 95% confidence interval, 0.77 to 0.86) and calibration (Hosmer–Lemeshow χ^2-statistic=6.46; *P*=0.60). A simple risk score sheet based on the cardiovascular risk prediction model was also presented. We developed and validated a new cardiovascular risk prediction model in a general population of Japanese. The risk prediction model would provide a useful guide to estimate absolute risk of cardiovascular disease and to treat individual risk factors.
Hypertension Research (2009) **32,** 1119–1122; doi:10.1038/hr.2009.161; published online 18 September 2009

Keywords: cardiovascular disease; epidemiology; risk factors; risk prediction model

INTRODUCTION

Cardiovascular disease is estimated to be one of the leading causes of death in Japan, as well as other countries around the world, placing a burden on the community.[1] Although the incidence and mortality of cardiovascular disease in Japan have declined over several decades, the risk of cardiovascular events remains high.[2] Additional protection will require an effective strategy for prevention of cardiovascular disease. Among a number of cardiovascular prevention strategies, high-risk approaches are likely to be one of the most effective strategies for prevention of cardiovascular disease.[3] To identify individuals at high risk of cardiovascular disease, a number of risk prediction tools have been developed.[4–15] However, currently available risk prediction tools of cardiovascular disease are derived mainly from studies carried out in Western populations and few risk prediction tools are developed for general Japanese populations. The objective of this paper is to develop a new cardiovascular risk prediction model and to validate its performance in a general population of Japanese.

METHODS

Study design and participants

Since 1961, we have been conducting a long-term prospective cohort study of cardiovascular disease in the town of Hisayama, a suburb of Fukuoka City in Southern Japan.[2,16,17] In 1988, a screening survey for this study was performed in the town. A total of 2742 residents aged 40 years or older (80.9% of the total population of this age group) consented to participate in the examination.[2,18–21] After the exclusion of 106 subjects with a history of cardiovascular disease and two subjects who died during the examination, the remaining 2634 individuals were enrolled in this study.

The ethics committee of Kyushu University approved this study, participants provided written informed consent, and the procedures followed were in accordance with national guidelines.

Follow-up survey

The subjects were followed up prospectively from December 1988 to November 2002 by repeated health examinations. A detailed description of the study methods has been published previously.[2,18–21] In brief, the health status of any subject who had not undergone a regular examination or who had moved out

[1]Department of Environmental Medicine, Graduate School of Medical Sciences, Kyushu University, Fukuoka, Japan and [2]Department of Medicine and Clinical Science, Graduate School of Medical Sciences, Kyushu University, Fukuoka, Japan
Correspondence: Dr H Arima, Department of Environmental Medicine, Graduate School of Medical Sciences, Kyushu University, 3-1-1 Maidashi, Higashi-ku, Fukuoka 812-8582, Japan.
E-mail: harima@envmed.med.kyushu-u.ac.jp
Received 29 June 2009; revised 13 August 2009; accepted 20 August 2009; published online 18 September 2009

of town was checked yearly by mail or telephone. We also established a daily monitoring system among the study team and local physicians or members of the town's Health and Welfare Office. When a subject died, an autopsy was performed at the Departments of Pathology of Kyushu University. During the follow-up period, 577 subjects died, of whom 438 (75.9%) underwent autopsy. Only one participant was lost to follow-up.

Outcomes

The primary outcome of the present analysis was cardiovascular disease. Cardiovascular disease was defined as first-ever development of coronary heart disease or stroke. The criteria for a diagnosis of coronary heart disease included first-ever acute myocardial infarction, silent myocardial infarction, sudden cardiac death within 1 h after the onset of acute illness, or coronary artery disease followed by coronary artery bypass surgery or angioplasty.[2] Acute myocardial infarction was diagnosed when a subject met at least two of the following criteria: (1) typical symptoms, including prolonged severe anterior chest pain; (2) abnormal cardiac enzymes more than twice the upper limit of the normal range; (3) evolving diagnostic electrocardiographic changes; and (4) morphological changes, including local asynergy of cardiac wall motion on echocardiography, persistent perfusion defect on cardiac scintigraphy, or myocardial necrosis or scars >1 cm long accompanied by coronary atherosclerosis at autopsy. Silent myocardial infarction was defined as myocardial scarring without any historical indication of clinical symptoms or abnormal cardiac enzyme changes, and was detected by electrocardiography, echocardiography, cardiac scintigraphy or autopsy. Stroke was defined as a sudden onset of nonconvulsive and focal neurological deficit persisting for >24 h. The diagnosis of stroke and the determination of its pathological type were based on the clinical history, neurological examination and all available clinical data, including brain CT/MRI and autopsy findings.[2]

Risk factors

Sitting blood pressure was measured three times at the right upper arm using a sphygmomanometer after 5 min of rest; an average of three measurements was used for the analysis. Plasma glucose levels were determined by the glucose-oxidase method, and diabetes was defined by a 75 g oral glucose tolerance test and by fasting ($\geqslant 7.0\,\mathrm{mmol\,l^{-1}}$) or postprandial ($\geqslant 11.1\,\mathrm{mmol\,l^{-1}}$) blood glucose levels or by the use of hypoglycemic agents. Total cholesterol, high-density lipoprotein cholesterol and triglyceride levels were determined enzymatically. Low-density lipoprotein (LDL) cholesterol level was estimated using the Friedewald formula.[22] Information on smoking habits was obtained using a standard questionnaire and was classified as either current or not.

Statistical analysis

Two-thirds of the study participants (n=1756) were randomly assigned to a risk prediction model derivation cohort and the remaining one-third (n=878) were reserved as an independent validation cohort using random digits generated by the Mersenne Twister method.[23] Among subjects allocated to the derivation cohort, a new risk prediction model was developed using Cox's proportional hazards model. Covariates included in Cox's proportional hazards model were age, sex, systolic blood pressure, diabetes, LDL cholesterol, high-density lipoprotein cholesterol and smoking habits that were traditional risk factors for cardiovascular disease established in the Hisayama study.[16,17,20,21] The performance of the risk prediction model was then tested among subjects allocated to the validation cohort. Ability of the risk prediction model to discriminate persons who experience a cardiovascular disease from those who do not were evaluated using c-statistic,[24] and calibration of the risk prediction model was evaluated using a Hosmer–Lemeshow χ^2-statistic with 8 d.f. The cardiovascular risk prediction model was translated into a risk score sheet using methods developed in the Framingham Heart Study.[25] To facilitate easier understanding of the concept of risk, 'vascular age' was also included in the risk score sheet. An individual's vascular age was calculated as the age of a person with the same predicted risk but with all other risk factor levels in optimal ranges.[10] All analyses were performed using the SAS software package (SAS Institute, Cary, NC, USA).

Table 1 Baseline characteristics in the derivation and the validation cohorts

	Derivation cohort (n=1756)	*Validation cohort (n=878)*
Age, years	59 (12)	59 (12)
Men	43%	40%
Systolic blood pressure, mm Hg	134 (21)	133 (22)
Diastolic blood pressure, mm Hg	78 (12)	77 (11)
Diabetes	11%	13%
LDL cholesterol, mg per 100 ml	131 (43)	133 (41)
HDL cholesterol, mg per 100 ml	50 (12)	50 (12)
Current smoker	24%	27%

Abbreviations: HDL, high-density lipoprotein; LDL, low-density lipoprotein.
Values are means (s.d.) or frequencies.
SI conversion factors: to convert LDL and HDL cholesterol to millimoles per liter, multiply by 0.0259.

Table 2 Regression coefficients and hazard ratios for the cardiovascular risk prediction model in the derivation cohort

	β	*Hazard ratio*	*95% CI*
Age, years	0.05775	1.059	1.046–1.073
Men	0.55569	1.743	1.264–2.404
Systolic blood pressure, mm Hg	0.01701	1.017	1.011–1.023
Diabetes	0.51977	1.682	1.193–2.370
LDL cholesterol, mg per 100 ml	0.00257	1.003	0.999–1.006
HDL cholesterol, mg per 100 ml	−0.01182	0.988	0.977–1.000
Current smoker	0.35287	1.423	1.024–1.978

Abbreviations: 95% CI, 95% confidence interval; HDL, high-density lipoprotein; LDL, low-density lipoprotein.
SI conversion factors: to convert LDL and HDL cholesterol to millimoles per liter, multiply by 0.0259.

RESULTS

The baseline characteristics of the subjects allocated to the derivation cohort and those to the validation cohort are shown in Table 1. There were no clear differences in these baseline characteristics between two cohorts.

During 14 years of follow-up, 216 cardiovascular events were observed in the derivation cohort and 125 in the validation cohort. The cardiovascular risk prediction model including covariates of age, sex, systolic blood pressure, diabetes, LDL cholesterol, high-density lipoprotein cholesterol and smoking habits were developed in the derivation cohort. The multivariate-adjusted regression coefficients and hazard ratios for the risk prediction model are shown in Table 2.

The performance of the risk prediction model was then evaluated among the validation cohort. In terms of discrimination, the c-statistic was as high as 0.81 (95% confidence interval, 0.77 to 0.86). Figure 1 demonstrates the calibration plots comparing actual and predicted cardiovascular events by deciles of risk. The calibration χ^2-statistic for the risk prediction model was 6.46 (d.f.=8), indicating excellent goodness of fit (P=0.60). The top 30% of predicted risk identified 70% of subjects who experienced cardiovascular disease during follow-up (sensitivity). Proportion of subjects without cardiovascular events who were not in the top 30% of predicted risk was 79% (specificity).

Tables 3 and 4 provide risk score sheets that can be used for estimation of the multivariable risk of cardiovascular disease at 10

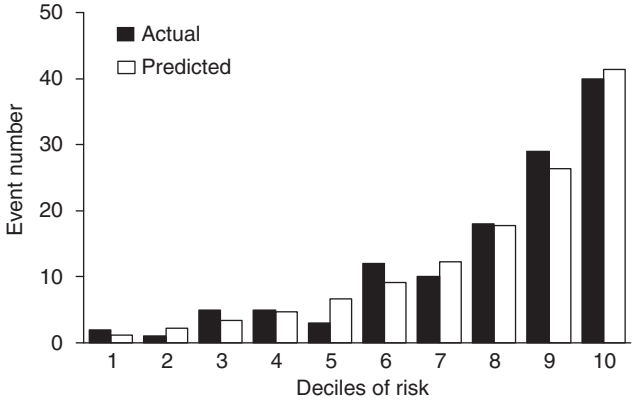

Figure 1 Actual and predicted cardiovascular events by deciles of risk in the validation cohort. Hosmer–Lemeshow χ^2-statistic=6.46, d.f.=8, P=0.60.

Table 3 Cardiovascular risk points

Points	*Age (years)*	*Sex*	*SBP (mm Hg)*	*Diabetic*	*LDL cholesterol (mg per 100 ml)*	*HDL cholesterol (mg per 100 ml)*	*Smoker*
0	40–44	Women	<119	No	<140	⩾40	No
1	45–49		120–139		⩾140	<40	Yes
2	50–54	Men	140–159	Yes			
3	55–59		160–179				
4	60–64		⩾180				
5	65–69						
6	70–74						
7	75–79						
8	⩾80						

Abbreviations: HDL, high-density lipoprotein; LDL, low-density lipoprotein; SBP, systolic blood pressure.
SI conversion factors: to convert LDL and HDL cholesterol to millimoles per liter, multiply by 0.0259.

Table 4 Estimated cardiovascular risk at 10 years and vascular age according to risk points

Points	*Risk[a] (%)*	*Vascular age for men[b] (years)*	*Vascular age for women[b] (years)*
0	1.4	—	40–44
1	1.8	—	45–49
2	2.4	40–44	50–54
3	3.2	45–49	55–59
4	4.2	50–54	60–64
5	5.6	55–59	65–69
6	7.4	60–64	70–74
7	9.8	65–69	75–79
8	12.8	70–74	80–84
9	16.7	75–79	85–89
10	21.7	80–84	90–94
11	27.8	85–89	95–99
⩾12	>30	⩾90	⩾100

[a]Estimated cardiovascular risk at 10 years.
[b]Age of a person with the same predicted risk but with all other risk factor levels in optimal ranges.

years. Table 4 also provide a different quantification of the same risk in the form of vascular age.

DISCUSSION

In this paper, a new risk prediction model of cardiovascular disease has been developed using data obtained from a prospective cohort study of a general Japanese population. The risk prediction model demonstrated good performance in regard to both discrimination and calibration. A simple risk score sheet based on the cardiovascular risk prediction model was also presented. This simple risk prediction tool of cardiovascular disease for Japanese would provide a useful guide to estimate absolute risk of cardiovascular disease and to treat individual risk factors.

Large-scale cohort studies have developed a number of risk prediction tools of cardiovascular disease.[4–15] However, these risk prediction tools were mainly derived from studies carried out in Western populations and few risk prediction tools are developed among general Japanese populations. The NIPPON DATA 80 derived a cardiovascular risk prediction tool, in which age, sex, systolic blood pressure, glucose levels, total cholesterol and smoking habits were included as risk factors, using data obtained from a 19-year prospective cohort study of general Japanese populations, although the outcome of NIPPON DATA 80 risk charts was death from cardiovascular causes.[5] The Jichi Medical School (JMS) cohort study developed 10-year risk prediction tools for incidence of myocardial infarction[14] and stroke,[15] in which age, sex, systolic blood pressure, diabetes, total cholesterol and smoking habits were included as risk factors, using data obtained from a population-based prospective study of general Japanese populations. The present analysis from the Hisayama study developed a new risk prediction tool for incidence of cardiovascular disease in a general population of Japanese using similar risk factors used in the previous observational studies of Japanese. Cumulative incidence rates of cardiovascular events at 10 years estimated from the present risk prediction tool were almost similar to combined risks of myocardial infarction and stroke obtained from the JMS risk charts[14,15] and this finding supports the validity and the generalizability of the Hisayama risk prediction model.

Several limitations of our study should be discussed. One limitation is a lack of external validation of the risk prediction model. However, split sample validation is an established method for internal validation of a risk prediction model and is widely used in other studies.[9,12] Similarity to the JMS risk chart[14,15] also supports the validity of the Hisayama risk prediction model. Another limitation is that LDL cholesterol, as a continuous variable, did not reach statistical significance in the derivation cohort. However, LDL cholesterol is an established risk factor for cardiovascular disease in the Hisayama study[21] and thus we included LDL cholesterol into the risk prediction model. A third limitation is that our findings are based on a one-time measurement of risk factors (for example, systolic blood pressure, plasma glucose levels, LDL cholesterol levels and high-density lipoprotein cholesterol levels), which may not accurately reflect the status of a study participant. A fourth limitation is that the value of LDL cholesterol was not directly assayed but was calculated by the Friedewald equation,[22] although the equation has been adopted in substantial epidemiologic and clinical studies of LDL cholesterol and cardiovascular disease. These limitations may have resulted in underestimation of the predicted risk among subjects at high risk of cardiovascular disease.

In conclusion, we developed and validated a new cardiovascular risk prediction model in a general population of Japanese. The risk prediction model would provide a useful guide to identify the individuals at high risk of cardiovascular disease in Japan. High-risk

approaches for the prevention of cardiovascular disease using the present risk prediction tool are likely to provide additional protection against the burden of cardiovascular disease in Japan.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

ACKNOWLEDGEMENTS

This study was supported in part by Grants-in-Aid for Scientific Research A (no.18209024) and C (no.19590633) from the Ministry of Education, Culture, Sports, Science, and Technology of Japan.

1 World Health Organization. *The atlas of heart disease and stroke*. World Health Organization: Geneva, 2004.
2 Kubo M, Kiyohara Y, Kato I, Tanizaki Y, Arima H, Tanaka K, Nakamura H, Okubo K, Iida M. Trends in the incidence, mortality, and survival rate of cardiovascular disease in a Japanese community: the Hisayama study. *Stroke* 2003; **34**: 2349–2354.
3 World Health Organization. *Preventing chronic disease: a vital investment*. World Health Organization: Geneva, 2005.
4 Conroy RM, Pyorala K, Fitzgerald AP, Sans S, Menotti A, De Backer G, De Bacquer D, Ducimetiere P, Jousilahti P, Keil U, Njolstad I, Oganov RG, Thomsen T, Tunstall-Pedoe H, Tverdal A, Wedel H, Whincup P, Wilhelmsen L, Graham IM, for the SCORE Project Group. Estimation of ten-year risk of fatal cardiovascular disease in Europe: the SCORE project. *Eur Heart J* 2003; **24**: 987–1003.
5 NIPPON DATA 80 Research Group. Risk assessment chart for death from cardiovascular disease based on a 19-year follow-up study of a Japanese representative population. *Circ J* 2006; **70**: 1249–1255.
6 Wu Y, Liu X, Li X, Li Y, Zhao L, Chen Z, Rao X, Zhou B, Detrano R, Liu K, for the USA-PRC Collaborative Study of Cardiovascular and Cardiopulmonary Epidemiology Research Group and the Chine Multicenter Collaborative Study of Cardiovascular Epidemiology (China MUCA) Research Group. Estimation of 10-year risk of fatal and nonfatal ischemic cardiovascular diseases in Chinese adults. *Circulation* 2006; **114**: 2217–2225.
7 Asia pacific Cohort Studies Collaboration. Cardiovascular risk prediction tools for populations in Asia. *J Epidemiol Community Health* 2007; **61**: 115–121.
8 Mendis S, Lindholm LH, Mancia G, Whitworth J, Alderman M, Lim S, Heagerty T. World Health Organization (WHO) and International Society of Hypertension (ISH) risk prediction charts: assessment of cardiovascular risk for prevention and control of cardiovascular disease in low and middle-income countries. *J Hypertens* 2007; **25**: 1578–1582.
9 Ridker PM, Buring JE, Rifai N, Cook NR. Development and validation of improved algorithms for the assessment of global cardiovascular risk in women: the Reynolds Risk Score. *JAMA* 2007; **297**: 611–619.
10 D'Agostino RB, Vasan RS, Pencina MJ, Wolf PA, Cobain M, Massaro JM, Kannel WB. General cardiovascular risk profile for use in primary care: the Framingham Heart Study. *Circulation* 2008; **117**: 743–753.
11 Gaziano TA, Young CR, Fitzmaurice G, Atwood S, Gaziano JM. Laboratory-based versus non-laboratory-based method for assessment of cardiovascular disease risk: the NHANES I Follow-up Study cohort. *Lancet* 2008; **371**: 923–931.
12 Hippisley-Cox J, Coupland C, Vinogradova Y, Robson J, Minhas R, Sheikh A, Brindle P. Predicting cardiovascular risk in England and Wales: prospective derivation and validation of QRISK2. *Br Med J* 2008; **336**: 1475–1482.
13 Teramoto T, Ohashi Y, Nakaya N, Yokoyama S, Mizuno K, Nakamura H, for the MEGA Study Group. Practical risk prediction tools for coronary heart disease in mild to moderate hypercholesterolemia in Japan: originated from the MEGA study data. *Circ J* 2008; **72**: 1569–1575.
14 Matsumoto M, Ishikawa S, Kayaba K, Gotoh T, Nago N, Tsutsumi A, Kajii E, the Jichi Medical School (JMS) Cohort Study Group. Risk charts illustrating the 10-year risk of myocardial infarction among residents of Japanese rural communities: the JMS Cohort Study. *J Epidemiol* 2009; **19**: 94–100.
15 Ishikawa S, Matsumoto M, Kayaba K, Gotoh T, Nago N, Tsutsumi A, Kajii E, the Jichi Medical School (JMS) Cohort Study Group. Risk charts illustrating the 10-year risk of stroke among residents of Japanese rural communities: the JMS Cohort Study. *J Epidemiol* 2009; **19**: 101–106.
16 Kiyohara Y, Ueda K, Fujishima M. Smoking and cardiovascular disease in the general population in Japan. *J Hypertens* 1990; **8**: S9–S15.
17 Arima H, Tanizaki Y, Kiyohara Y, Tsuchihashi T, Kato I, Kubo M, Tanaka K, Ohkubo K, Nakamura H, Abe I, Fujishima M, Iida M. Validity of the JNC VI recommendations for the management of hypertension in a general population of Japanese elderly: the Hisayama Study. *Arch Intern Med* 2003; **163**: 361–366.
18 Arima H, Kiyohara Y, Kato I, Tanizaki Y, Kubo M, Iwamoto H, Tanaka K, Abe I, Fujishima M. Alcohol reduces insulin-hypertension relationship in a general population: the Hisayama study. *J Clin Epidemiol* 2002; **55**: 863–869.
19 Arima H, Kubo M, Yonemoto K, Doi Y, Ninomiya T, Tanizaki Y, Hata J, Matsumura K, Iida M, Kiyohara Y. High-sensitivity C-reactive protein and coronary heart disease in a general population of Japanese: the Hisayama study. *Arterioscler Thromb Vasc Biol* 2008; **28**: 1385–1391.
20 Doi Y, Kubo M, Yonemoto K, Ninomiya T, Iwase M, Arima H, Hata J, Tanizaki Y, Iida M, Kiyohara Y. Fasting plasma glucose cutoff for diagnosis of diabetes in a Japanese population. *J Clin Endocrinol Metab* 2008; **93**: 3425–3429.
21 Imamura T, Doi Y, Arima H, Yonemoto K, Hata J, Kubo M, Tanizaki Y, Ibayashi S, Iida M, Kiyohara Y. LDL cholesterol and the development of stroke subtypes and coronary heart disease in a general Japanese population: the Hisayama study. *Stroke* 2009; **40**: 382–388.
22 Friedewald W, Levy R, Fredrickson D. Estimation of the concentration of low density lipoprotein cholesterol in plasma without the use of the preparative ultracentrifuge. *Clin Chem* 1972; **18**: 499–502.
23 Matsumoto M, Nishimura T. Mersenne Twister: a 623-dimensionally equidistributed uniform pseudorandom number generator. *ACM Trans on Modeling and Computers Simulations* 1998; **8**: 3–30.
24 Pencina MJ, D'Agostino RB. Overall C as a measure of discrimination in survival analysis: model specific population value and confidence interval estimation. *Stat Med* 2004; **23**: 2109–2123.
25 Sullivan LM, Massaro JM, D'Agostino RB. Presentation of multivariate data for clinical use: the Framingham study risk score functions. *Stat Med* 2004; **23**: 1631–1660.